New Results on Normality

David H. Bailey
Lawrence Berkeley National Laboratory
Berkeley, CA, USA 94720
dhbailey@lbl.gov

Collaborators:
Jonathan M. Borwein, Simon Fraser University, Canada
Richard E. Crandall, Center for Advanced Computation, Reed College
Carl Pomerance, AT&T Bell Laboratories

Normality

The real number α is *normal* to base b if every sequence of m digits in the base-b expansion of α appears with limiting frequency b^{-m}.

Almost all real numbers are normal (from measure theory). Widely believed to be normal base b for all bases b:

- π and e.
- $\log 2$ and $\sqrt{2}$.
- The golden mean $\tau = (1 + \sqrt{5})/2$.
- *Every* irrational algebraic number.
- Many other "natural" irrational constants.

But there are *no* proofs for any of these constants, for any base. Normality proofs exist only for handful of artifically constructed constants, such as Champernowne's number: 0.1234567891011121314...

Peter Borwein's Observation on the Binary Digits of log 2

In 1995, Peter Borwein observed that a segment of binary digits of $\log 2$ beginning after the first d bits can be calculated by using a very simple algorithm:

Let $\{\cdot\}$ denote the fractional part. Then we can write

$$
\begin{aligned}
\{2^d \log 2\} &= \left\{2^d \sum_{k=1}^{\infty} \frac{1}{k 2^k}\right\} = \left\{\sum_{k=1}^{\infty} \frac{2^{d-k}}{k}\right\} \\
&= \left\{\left\{\sum_{k=1}^{d} \frac{2^{d-k}}{k}\right\} + \sum_{k=d+1}^{\infty} \frac{2^{d-k}}{k}\right\} \\
&= \left\{\left\{\sum_{k=1}^{d} \frac{2^{d-k} \bmod k}{k}\right\} + \sum_{k=d+1}^{\infty} \frac{2^{d-k}}{k}\right\}
\end{aligned}
$$

- The numerators $2^{d-k} \bmod k$ can be very rapidly evaluated using the binary algorithm for exponentiation performed modulo k.
- Only a few terms of the second summation need be evaluated.
- All computations can be done with ordinary 64-bit floating-point arithmetic.

The BBP Formula for π

By applying my PSLQ computer program to a set of computed constants for which formulas of this type were known, with the numerical value of π appended, Peter Borwein and Simon Plouffe found this formula for π:

$$\pi = \sum_{k=0}^{\infty} \frac{1}{16^k} \left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6} \right)$$

Thus segments of base-16 (or base-2) digits π beginning at arbitrary positions can be rapidly calculated, as with $\log 2$.

Since 1996, BBP-type formulas have been discovered for numerous other constants.

Question: Why wasn't this formula discovered 250 years ago?

A Connection Between BBP-Type Formulas and Normality

Theorem: The BBP-type constant

$$\alpha = \sum_{k=0}^{\infty} \frac{p(k)}{b^k q(k)}$$

(where $p(k)$ and $q(k)$ are integer polynomials, $\deg p < \deg q$ and q has no zeroes for positive k) is normal base b if and only if the sequence $x_0 = 0$, and

$$x_n = \left\{ b x_{n-1} + \frac{p(n)}{q(n)} \right\}$$

is equidistributed in the unit interval.

Proof Sketch: Let α_n be the base-b expansion of α after the n-th digit. Following the BBP approach, we can write

$$\begin{aligned} \alpha_n &= \left\{ \sum_{k=0}^{n} \frac{b^{n-k} p(k)}{q(k)} \right\} + \sum_{k=n+1}^{\infty} \frac{b^{n-k} p(k)}{q(k)} \\ &= \left\{ b \alpha_{n-1} + \frac{p(n)}{q(n)} \right\} + E_n \end{aligned}$$

where E_n goes to zero.

Two Examples

1. Let $x_0 = 0$, and

$$x_n = \left\{ 2x_{n-1} + \frac{1}{n} \right\}$$

Is (x_n) equidistributed in $[0, 1)$?

2. Let $x_0 = 0$ and

$$x_n = \left\{ 16x_{n-1} + \frac{120n^2 - 89n + 16}{512n^4 - 1024n^3 + 712n^2 - 206n + 21} \right\}$$

Is (x_n) equidistributed in $[0, 1)$?

If answer to Question 1 is “yes”, then $\log 2$ is normal to base 2.

If answer to Question 2 is “yes”, then π is normal to base 16 (and hence to base 2 also).

A Class of Provably Normal Constants

Using the BBP approach, Richard Crandall and I have now proven normality for a class of constants, the simplest instance of which is

$$\begin{aligned} \alpha_{2,3} &= \sum_{k=1}^{\infty} \frac{1}{3^k 2^{3^k}} \\ &= 0.041883680831502985071252898624571682426096\ldots_{10} \\ &= \mathtt{0.0AB8E38F684BDA12F684BF35BA781948B0FCD6E9E0}\ldots_{16}. \end{aligned}$$

$\alpha_{2,3}$ was actually proven normal base 2 in a little-known paper by Stoneham in 1977. Crandall and I proved normality and transcendence for an uncountably infinite class that includes $\alpha_{2,3}$:

$$\alpha_{2,3}(r) = \sum_{k=1}^{\infty} \frac{1}{3^k 2^{3^k + r_k}}$$

where r_k is the k-th bit in the binary expansion of $r \in (0, 1)$.

These constants also possess the rapid individual digit computation property. The googol-th binary digit of $\alpha_{2,3}$ is zero.

Probability Measures and the Birkoff Ergodic Theorem

Definition. Given a probability measure μ on a measure space Ω, the transformation T is said to be *ergodic* if: (1) for every measurable set A, $\mu(T^{-1}A) = \mu(A)$, and (2) if $T^{-1}A = A$ then $\mu(A) = 0$ or 1.

Example: $\Omega = [0, 1)$ is the unit circle mod 1, μ is ordinary Lebesgue measure, and $T(x) = \{2x\}$, where $\{\cdot\}$ denotes fractional part.

Ergodic Theorem. Let $f(t)$ be an integrable function on a measure space with probability measure μ, and let T be an ergodic transformation. Then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) = \int f \, d\mu \quad \text{for a.e. } x(\mu)$$

where "for a.e. $x(\mu)$" means for all x except for a set N with $\mu(N) = 0$.

The ergodic theorem can be thought of as the law of large numbers extended to a general measure space.

Equivalence of Absolutely Continuous Measures

Lemma. Let μ be a probability measure and T an ergodic transformation. Suppose that ν is another measure for which T is ergodic, and further ν is absolutely continuous with respect to μ (i.e., $\nu(A) = 0$ if and only if $\mu(A) = 0$). Then $\mu = \nu$.

Proof. Applying ergodic theorem to $f(t) = I_A(t)$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^n x) \;=\; \int f(t)\, d\mu(t) \;=\mu(A) \quad \text{for a.e. } x(\mu).$$

Since ν is absolutely continuous with respect to μ, the above holds a.e. $x(\nu)$ as well. Now since T preserves the measure ν, we can write, for $n > 0$,

$$\begin{aligned} \nu(A) \;&=\; \int f(t)\, d\nu(t) \;=\; \frac{1}{n} \sum_{i=0}^{n-1} \int f(T^i x)\, d\nu(x) \\ &=\; \int \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x)\, d\nu(x) \;\to\; \int \mu(A) d\nu \;=\; \mu(A) \end{aligned}$$

by the dominated convergence theorem.

The Hot Spot Lemma

Lemma. The real constant α is normal base b if and only if there exists a constant C such that for every subinterval $[c, d) \subset [0, 1)$,

$$\limsup_{n \geq 1} \frac{\#_{0 \leq j < n}(\{b^j \alpha\} \in [c, d))}{n} \leq C(d - c).$$

Proof. Let μ denote ordinary Lebesgue measure on $[0, 1)$ (the unit interval mod 1), let $T(x) = \{2x\}$, and let ν be the measure on $[0, 1)$, defined on the interval $[c, d)$ to be the LHS of the condition in the hot spot lemma. It is easily seen that T preserves both μ and ν, so that T is ergodic for both μ and ν.

The condition in the hot spot lemma is easily seen to imply that ν is absolutely continuous with respect to μ. Thus by previous lemma, $\mu = \nu$, or in other words $\{b^k \alpha\}$ is uniformly distributed in the unit interval mod 1. This implies that α is normal base b.

The BBP Sequence Associated with $\alpha_{2,3}$.

The BBP sequence for

$$\alpha_{2,3} = \sum_{k=1}^{\infty} \frac{1}{3^k 2^{3^k}}$$

is: $x_0 = 0$, and $x_n = \{2x_{n-1} + r_n\}$, where $r_n = 1/n$ if $n = 3^k$, but zero otherwise. The sequence (x_n) is easily seen to be the concatenation of primitive linear congruential pseudorandom sequences, each of length $2 \cdot 3^k$:

$0,$ repeated 3 times,

$\frac{1}{3}, \frac{2}{3},$ repeated 3 times,

$\frac{4}{9}, \frac{8}{9}, \frac{7}{9}, \frac{5}{9}, \frac{1}{9}, \frac{2}{9},$ repeated 3 times,

$\frac{13}{27}, \frac{26}{27}, \frac{25}{27}, \frac{23}{27}, \frac{19}{27}, \frac{11}{27}, \frac{22}{27}, \frac{17}{27}, \frac{7}{27}, \frac{14}{27}, \frac{1}{27}, \frac{2}{27}, \frac{4}{27}, \frac{8}{27}, \frac{16}{27}, \frac{5}{27}, \frac{10}{27}, \frac{20}{27},$

repeated 3 times, etc.

Simple Proof That $\alpha_{2,3}$ Is Normal Base 2

Proof. Note that for $n < 3^{k+1}$, each x_n is a multiple of $1/3^k$, and each fraction $j/3^k$, $0 \le j < 3^k$ appears exactly three times in the sequence. Also note that

$$|x_n - \alpha_n| = \left| \sum_{k=n+1}^{\infty} 2^{n-k} r_k \right| < \frac{1}{2n}$$

where $\alpha_n = \{2^n \alpha_{2,3}\}$. Given n, let m be the largest power of 3 less than n, and assume that n is large enough so that $n > m > 1/(d-c)$. Now note that the interval $[c - 1/(2n), d + 1/(2n))$ contains exactly $m(d-c)$ (or possibly one more) multiples of $1/m$, and thus can contain at most three times this many occurrences of x_j in the first n elements. Thus we can write

$$\begin{aligned} \frac{\#_{0 \le j < n}(\alpha_j \in [c,d))}{n(d-c)} &\le \frac{\#_{0 \le j < n}(x_j \in [c - 1/(2n), d + 1/(2n)))}{n(d-c)} \\ &\le \frac{3[m(d-c)+1]}{n(d-c)} < \frac{3[m(d-c)+1]}{m(d-c)} \\ &= 3 + \frac{3}{m(d-c)} < 6 \end{aligned}$$

Thus $\alpha_{2,3}$ is normal base 2, by the hot spot lemma.

A Result for Irrational Square Roots

Lemma. Let $B(a)$ denote the number of one bits in the integer a. Then for any two positive integers a and b, $B(a+b) \leq B(a)+B(b)$, and $B(ab) \leq B(a)B(b)$. Proof is by induction on number of bits in a and b.

Theorem: Let $B_n(\alpha)$ be the number of one bits in the first n bits of α. If α is the square root of an integer or rational, then for some constant C,

$$\liminf_{n\to\infty} \frac{B_n(\alpha)}{C\sqrt{n}} \geq 1$$

Proof Sketch. Consider $b_{50} = \sqrt{2}$ truncated to 50 bits:

$$\begin{aligned} b_{50} &= 1.01101010000010011110011001100111111100111011110011 \\ b_{50}^2 &= 1.111111111111111111111111111111111111111111111111110011001\ldots \end{aligned}$$

Note that the expansion of b_{50}^2 is all ones, up to approximately 50 bits. Thus we conclude that the first n bits of $\sqrt{2}$ must have at least $\sqrt{n}$ ones, or else the product won't have enough ones to fill the first n bits.

This observation leads to a rigorous proof of Theorem A.

A Result for General Irrational Algebraic Numbers

Theorem. For any real algebraic irrational α, we have

$$\liminf_{n\to\infty} \frac{B(n)}{\log_2 n} \geq 1 \tag{1}$$

Proof. Let p_n denote the position of the n-th one in the binary expansion of α. Note that we can write

$$\alpha = \sum_{k=1}^{p_n} 2^{-k} + \sum_{k=p_{n+1}}^{\infty} 2^{-k}$$

Write the first term as the fraction C_n/D_n in lowest terms, and note $D_n = 2^{p_n}$. According to Roth's theorem, since α is algebraic, then given $\epsilon > 0$, there is some N such that for all $n > N$ and for any E,

$$\left|\alpha - \frac{C_n}{D_n}\right| < \frac{E}{D_n^2}$$

only finitely often. Thus for every ϵ, there is an N such that for every $n > N$, we have $p_{n+1} < (2+\epsilon)p_n$. It follows that $p_n < p_N(2+\epsilon)^{n-N} < K(2+\epsilon)^n$, and the result follows with some additional effort.

A Stronger Result for General Irrational Algebraic Numbers

Theorem. Let α be an irrational algebraic number of degree $d \geq 2$, and let a_d be the leading (high-order) coefficient for the minimal polynomial satisfied by α. Then for any $\epsilon > 0$,

$$B_n(\alpha) > (1-\epsilon)a_d^{1/d} n^{1/d}$$

for all sufficiently large n.

The proof is given in a new manuscript (March 2003) by Jonathan Borwein, Richard Crandall, Carl Pomerance and myself.

Corollary. Let α be any positive real, and let F_n denote the Fibonacci numbers. Then these constants are transcendental:

$$\beta_1 = \sum_{n=1}^{\infty} \frac{1}{2^{\lfloor \alpha^n \rfloor}}$$

$$\beta_2 = \sum_{n=1}^{\infty} \frac{1}{2^{F_n}}$$

Additional result. This constant is not a quadratic irrational:

$$\beta_3 = \sum_{n=1}^{\infty} \frac{1}{2^{n^2}}$$

For Full Details

- David H. Bailey, Peter B. Borwein and Simon Plouffe, "On The Rapid Computation of Various Polylogarithmic Constants," *Mathematics of Computation*, vol. 66, no. 218, 1997, pp. 903–913.
- David H. Bailey, "A Compendium of BBP-Type Formulas," 2002.
- David H. Bailey and Richard E. Crandall, "On the Random Character of Fundamental Constant Expansions," *Experimental Mathematics*, June 2001.
- David H. Bailey and Richard E. Crandall, "Random Generators and Normal Numbers," *Experimental Mathematics*, to appear.
- David H. Bailey, Jonathan M. Borwein, Richard E. Crandall and Carl Pomerance, "On the binary expansions of algebraic numbers," March 2003.

These are available at:
`http://www.nersc.gov/~dhbailey/dhbpapers`